



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
13/492,540	06/08/2012	Charles Corfield	88999-8028.US00	3289

98783 7590 02/01/2017
Perkins Coie LLP - DEN General
P.O. Box 1247
Seattle, WA 98111-1247

EXAMINER

NGUYEN, TIMOTHY

ART UNIT	PAPER NUMBER
----------	--------------

2657

NOTIFICATION DATE	DELIVERY MODE
-------------------	---------------

02/01/2017

ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentprocurement@perkinscoie.com

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

Ex parte CHARLES CORFIELD and BRIAN MARQUETTE

Appeal 2016-002487
Application 13/492,540¹
Technology Center 2600

Before DEBRA K. STEPHENS, KARA L. SZPONDOWSKI, and
SHARON FENICK, *Administrative Patent Judges*.

FENICK, *Administrative Patent Judge*.

DECISION ON APPEAL

Appellants seek our review under 35 U.S.C. § 134(a) of the Examiner's Final Rejection of claims 1–17, all the pending claims in the present application. (Appeal Br. 2.) We have jurisdiction over the appeal under 35 U.S.C. § 6(b)(1).

We REVERSE.

Invention

Appellants' invention relates to a speech recognition system using two engines or modes: a first speech recognition engine or mode that is not a

¹ According to Appellants, the real party in interest is NVOQ INCORPORATED. (Appeal Br. 2.)

continuous speech recognition engine or mode, and a second speech recognition engine or mode for continuous speech recognition. (Spec. Abstract, ¶¶ 20–22.) Initially, the first speech recognition engine or mode is used to convert a user’s spoken input into a transcript. (*Id.* ¶¶ 26–27.) The input and the transcript are used as input to train the second speech recognition engine or mode. (*Id.* Abstract, ¶¶ 22–25.) Eventually, the second speech recognition engine or mode is used for transcription. (*Id.* Abstract, ¶ 27.)

Illustrative Claim

Claim 1, reproduced below with emphasis added, is illustrative:

1. A method performed on at least one processor for training a user profile for a continuous speech recognition engine, the method comprising the steps of:

providing an audio file wherein the audio file contains a pronunciation of a client;

recognizing the audio file as text data using a first speech recognition engine, *wherein the first speech recognition engine is not a continuous speech recognition engine*;

linking the audio file and the text data generated by the first speech recognition engine;

initially training a user profile of a second speech recognition engine, which is the continuous speech recognition engine, using the linked audio file and text data, wherein the user profile of the continuous speech recognition engine is initially trained using the linked audio and text data recognized by the first speech recognition engine and wherein only the first speech recognition engine recognizes audio until the user profile is at least initially trained and wherein the second speech recognition engine only recognizes audio after at least a portion of the initial training of the user profile is completed such that until at least the portion of the initial training of the user profile is complete only the first speech recognition engine recognizes audio.

Rejections

Appellants appeal the following rejections:

Claims 1–8 and 10–17 are rejected under pre-AIA 35 U.S.C. § 103(a) as unpatentable over Kahn (US 2006/0167686 A1; pub. Jul. 27, 2006) and Di Fabrizio et al. (US 7,869,998 B1; iss. Jan. 11, 2011) (hereinafter “Di Fabrizio”). (Final Action 7–13.)

Claim 9 is rejected under pre-AIA 35 U.S.C. § 103(a) as unpatentable over Kahn, Di Fabrizio, and Bantz et al. (US 2002/0169606 A1; pub. Nov. 14, 2002). (Final Action 14.)

ANALYSIS

The Examiner finds that the combination of Kahn and Di Fabrizio teaches or suggests the limitations of claim 1, including the limitation that a “*first speech recognition engine is not a continuous speech recognition engine*” and the limitation of “*initially training a user profile of a second speech recognition engine, which is the continuous speech recognition engine, using the linked audio file and text data . . . wherein only the first speech recognition engine recognizes audio until the user profile is at least initially trained and wherein the second speech recognition engine only recognizes audio after at least a portion of the initial training of the user profile is completed.*” (Final Action 7–9.) The Examiner finds that Kahn discloses the use of a first speech recognition engine (element 211), which is not a continuous speech recognition engine. (*Id.* at 8, citing Kahn Fig. 2, ¶¶ 6, 75.) The Examiner further finds that Kahn describes the training (“enrollment”) of speech recognition engines, and thus teaches or suggests that Kahn’s disclosed second speech recognition engine (element 213), which is a continuous speech recognition engine, would be trained on linked

audio file and text data obtained from the first speech recognition engine. (*Id.* at 8–9.) The Examiner acknowledges that Kahn does not teach that only the first speech recognition engine recognizes audio until the user profile is at least initially trained, but finds this to be taught or suggested by Di Fabrizio in combination with Kahn’s disclosure. (*Id.* at 9.)

Kahn relates to a method for creating a transcription of an audio file. (Kahn, Abstract.) Kahn provides one or more speech engines which each produce a transcription of the text. (*Id.* ¶¶ 74–75.) When a user corrects any errors in the transcription, a verbatim text is produced which may be used to train the speech engine. (*Id.* ¶¶ 102, 114–115.) The verbatim text may be used for iterative speech engine training of the speech engines. (*Id.* ¶ 102.)

Di Fabrizio relates to a voice-enabled help desk service, which uses a general-purpose model and a domain-specific model, to adapt the speech recognition module after deployment. (Di Fabrizio, Abstract, 2:54–63.) The help desk service is adapted as task-specific data is created through the use of the service. (*Id.* at 5:8–14.)

Appellants argue that Kahn does not teach or suggest a first speech recognition engine which is not a continuous language speech recognition engine. (Appeal Br. 7–9.) Appellants argue that the Examiner’s citation of paragraph 6 of Kahn to show that the first speech recognition engine is not a continuous speech recognition program is inapposite, because “this section of Kahn shows a conventional continuous speech recognition program,” and Appellants further argue that Kahn’s disclosure of “a sophisticated process of cross correction” relates to the use of “separate texts generated by each of the two continuous speech recognition engines.” (Appeal Br. 9.)

We agree with the Appellants. The Examiner finds that the process of enrollment, described in Kahn with respect to the conventional training of continuous speech recognition programs, teaches speech recognition which is “not natural language/not continuous.” (Final Action 8, citing Kahn ¶ 6.) Whether or not Kahn describes the possibility of enrollment using recognition which is not continuous, the enrollment the Examiner cites is described as occurring for “[c]onventional continuous speech recognition programs.” (Kahn ¶ 6.) Appellants are correct that “Kahn specifically describes a speech recognition system that *avoids* using the enrollment process.” (Appeal Br. 9.) Kahn specifically discloses that the Kahn invention is used to avoid enrollment: “the inventors have discovered that iteratively processing an audio file with a non-enrolled user through the correction session of the invention surprisingly increased the accuracy of the speech engine.” (Kahn ¶ 66.) Thus, even if the enrollment cited by the Examiner for continuous speech recognition programs taught or suggested such programs to acting for an initial period in a non-continuous fashion, the teachings or suggestions of the cited portions of the Kahn invention relating to recognition of audio by a first speech engine (e.g., to Kahn’s audio file and transcribed text “A,” and to Figure 2 and paragraph 75 of Kahn) (Final Action 8) relate only to continuous speech recognition.

Because we agree with at least one of the arguments advanced by Appellants, we need not reach the merits of Appellants’ other arguments. Accordingly, we do not sustain the Examiner’s obviousness rejection of independent claim 1. Additionally, we do not sustain the rejections of independent claim 11, or dependent claims 2–10 and 12–17, all argued in whole or in part on the same basis. (Appeal Br. 11, 13.)

DECISION

We reverse the Examiner's § 103(a) rejections of claims 1–17 as obvious.

REVERSED